

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions and listings of claims in the application:

LISTING OF CLAIMS:

1. (canceled)
2. (previously presented): The method of claim 5, wherein the structural representation of discourse is determined based on a theory of discourse analysis.
3. (original): The method of claim 2, wherein the theory of discourse analysis is at least one of the Linguistic Discourse Model, the Unified Linguistic Discourse Model, Rhetorical Structure Theory, Discourse Structure Theory and Structured Discourse Representation Theory.
4. (previously presented): The method of claim 5, wherein non-structural measures of relevance are determined based on at least one of statistics, keywords, knowledge bases.
5. (currently amended): A method of determining a hybrid text summary by a hybrid summarization system having a processor, a relevance score determination module, a structural representation of the discourse determination module and a percolation module, the method comprising the steps of:
determining discourse constituents for a text by the processor;

determining a structural representation of discourse for the text by the structural representation of the discourse determination module;

determining relevance scores for discourse constituents based on at least one non-structural measure of relevance by the relevance score determination module;

percolating relevance scores based on the structural representation of discourse by the percolation module; and

determining a hybrid text summary, by the processor, based on discourse constituents with relevance scores compared to a threshold relevance score,

wherein percolating the relevance scores comprises the steps of:

for each child discourse constituent node in the structural representation, assigning the relevance score of the child discourse constituent node to the parent discourse constituent node if the child discourse constituent node is more relevant;

for any subordinating nodes, assigning the relevance scores of the subordinated discourse constituent to the subordinating discourse constituent if the subordinated discourse constituent is more relevant; and

for any coordination nodes, assigning the relevance score of the most relevant child to other child discourse constituent nodes.

6. (currently amended): A method of determining a hybrid text summary by a hybrid summarization system having a processor, a relevance score determination module, a structural representation of the discourse determination module and a percolation module, the method comprising the steps of:

determining discourse constituents for a text by the processor;

determining a structural representation of discourse for the text by the structural representation of the discourse determination module;

determining relevance scores for discourse constituents based on at least one non-structural measure of relevance by the relevance score determination module;

percolating relevance scores based on the structural representation of discourse by the percolation module; and

determining a hybrid text summary, by the processor, based on discourse constituents with relevance scores compared to a threshold relevance score,

wherein percolating the relevance scores comprises the steps of:

for each child discourse constituent node in the structural representation, assigning the relevance score of the child discourse constituent node to the parent discourse constituent node if the child discourse constituent node is more relevant than its parent;

for each coordinated discourse constituent node, assigning the relevance score of the coordinated discourse constituent node to each preceding less relevant sibling node; for each child discourse constituent node that is not a coordinated discourse constituent node and is not a subordinated discourse constituent node, assigning the relevance score of the parent discourse constituent node to the child discourse constituent node if the parent discourse constituent is more relevant than the child;

for each coordinated discourse constituent node, assigning the relevance score of the parent discourse constituent node to the coordinated discourse constituent node, if the coordinated discourse node and all its siblings are less relevant than the parent node;

for each subordinated discourse constituent node, assigning the relevance score of the subordinated discourse constituent node to the subordinating discourse constituent if the subordinated discourse constituent is more relevant than the subordinating node; and

for each node, repeating these steps, until no node can be found whose relevance score is changed to the relevance score of another node.

7. (original): The method of claim 6, wherein the percolation of relevance scores is applied to progressively larger sets of linked nodes.

8. (currently amended): A method of determining a hybrid text summary by a hybrid summarization system having a processor, a relevance score determination module, a structural representation of the discourse determination module and a percolation module, the method comprising the steps of:

determining discourse constituents for a text by the processor;

determining a structural representation of discourse for the text by the structural representation of the discourse determination module;

determining relevance scores for discourse constituents by the relevance score determination module;

percolating relevance scores by the percolation module based on the structural representation of discourse comprising the steps of:

for each discourse constituent leaf node, determining the number of subordinated edges plus one;

determining a score based on the inverse of the number of subordinated edges +1;
for each discourse constituent node, assigning the score of a child discourse constituent node to the parent discourse constituent node, if the score is less relevant;
for any subordination discourse constituent node, assigning the score of the subordinated discourse constituent node to the subordinating discourse constituent node if the subordinated discourse constituent score is lower;
assigning the relevance scores of any coordination discourse constituent node to each child discourse constituent of the coordination if it is lower;
determining an adjusted relevance score based on the score and the subordination level; and
determining a hybrid text summary, by the processor, based on discourse constituents with relevance scores compared to a threshold relevance score.

9. (previously presented): The method of claim 5, further comprising the steps of:
determining every leaf discourse constituent containing an anaphor;
for each anaphor, determine any unique antecedent referents for the anaphor;
substituting the unique antecedent referent into the leaf discourse constituent for the anaphor;
removing the discourse constituent containing the unique antecedent referent from the set of the discourse constituents with relevance scores more relevant than the threshold relevance score.

10-11. (canceled)

12. (previously presented): The system of claim 15, wherein the structural representation of discourse is determined based on a theory of discourse analysis.

13. (original): The system of claim 12, wherein the theory of discourse analysis is at least one of the Linguistic Discourse Model, the Unified Linguistic Discourse Model, Rhetorical Structure Theory, Discourse Structure Theory and Structured Discourse Representation Theory.

14. (previously presented): The system of claim 15, wherein non-structural measures of relevance are determined based on at least one of: statistics, keywords, knowledge bases.

15. (previously presented): A system for determining hybrid text summaries comprising:

an input/output circuit for retrieving a text;

a processor for determining discourse constituents for the text and attaching the discourse constituents into a structural representation of discourse;

a relevance score determination circuit for determining relevance scores for the discourse constituents based on at least one non-structural measure of relevance; and

a percolation circuit for percolating discourse constituent relevance scores based on the structural representation of discourse and where the processor determines a hybrid text summary based on the discourse constituents with relevance scores exceeding a threshold relevance score,

wherein, for each child discourse constituent node in the structural representation, the percolation circuit assigns the relevance score of the child discourse constituent node to the parent discourse constituent node if the child discourse constituent node is more relevant;

for any subordinating nodes, the percolation circuit assigns the relevance scores of the subordinated discourse constituent to the subordinating discourse constituent if the subordinated discourse constituent is more relevant; and

for any coordination nodes, the percolation circuit assigns the relevance score of the most relevant child to other child discourse constituent nodes.

16. (previously presented): A system for determining hybrid text summaries comprising:

an input/output circuit for retrieving a text;

a processor for determining discourse constituents for the text and attaching the discourse constituents into a structural representation of discourse;

a relevance score determination circuit for determining relevance scores for the discourse constituents based on at least one non-structural measure of relevance; and

a percolation circuit for percolating discourse constituent relevance scores based on the structural representation of discourse and where the processor determines a hybrid text summary based on the discourse constituents with relevance scores exceeding a threshold relevance score,

wherein for each child discourse constituent node in the structural representation, the percolation circuit assigns the relevance score of the child discourse constituent node to the

parent discourse constituent node if the child discourse constituent node is more relevant than its parent;

for each coordinated discourse constituent node, the percolation circuit assigns the relevance score of the coordinated discourse constituent node to each preceding less relevant sibling node;

for each child discourse constituent node that is not a coordinated discourse constituent node and is not a subordinated discourse constituent node, the percolation circuit assigns the relevance score of the parent discourse constituent node to the child discourse constituent node if the parent discourse constituent is more relevant than the child;

for each coordinated discourse constituent node, the percolation circuit assigns the relevance score of the parent discourse constituent node to the coordinated discourse constituent node, if the coordinated discourse node and all its siblings are less relevant than the parent node;

for each subordinated discourse constituent node, the percolation circuit assigns the relevance score of the subordinated discourse constituent node to the subordinating discourse constituent if the subordinated discourse constituent is more relevant than the subordinating node; and

for each node, repeating these steps, until the percolation circuit can find no node whose relevance score is changed to the relevance score of another node.

17. (original): The system of claim 16, wherein the percolation is applied to progressively larger sets of linked nodes.

18. (previously presented): A system for determining hybrid text summaries comprising:

an input/output circuit for retrieving a text;

a processor for determining discourse constituents for the text and attaching the discourse constituents into a structural representation of discourse;

a relevance score determination circuit for determining relevance scores for the discourse constituents based on at least one non-structural measure of relevance;

a percolation circuit for percolating discourse constituent relevance scores based on the structural representation of discourse; wherein for each discourse constituent leaf node, the percolation circuit determines the number of subordinated edges plus one and a score based on the inverse of the number of subordinated edges +1;

for each discourse constituent node, the percolation circuit assigns the score of a child discourse constituent node to the parent discourse constituent, if the score is less relevant;

for any subordination discourse constituent node, the percolation circuit assigns the score of the subordinated discourse constituent node to the subordinating discourse constituent node if the subordinated discourse constituent score is lower;

the percolation circuit assigns the scores of any coordination discourse constituent node to each child discourse constituent of the coordination if it is lower; and

the processor determines an adjusted relevance score based on the score and the subordination level; and

a hybrid text summary based on the discourse constituents with relevance scores exceeding a threshold relevance score.

19. (previously presented): The system of claim 15, wherein the processor determines every leaf discourse constituent containing an anaphor;

for each anaphor, the processor determines any unique preceding referents for the anaphor;

the processor substitutes the unique antecedent referent into the preceding discourse constituent for the anaphor referent; and

the processor removes the preceding discourse containing the unique referent from the discourse constituents with relevance scores exceeding the threshold relevance score.

20. (previously presented): The system of claim 15, the percolation circuit determines every leaf discourse constituent containing an anaphor;

for each anaphor, the percolation circuit determines any unique preceding referents for the anaphor;

the percolation circuit substitutes the unique antecedent referent into the leaf discourse constituent for the anaphor;

the percolation circuit removes the discourse constituent containing the unique antecedent referent from the set of the discourse constituents with more relevant relevance scores.

21. (previously presented): The system of claim 15, wherein the processor determines important discourse constituent nodes based on a non-structural measure of relevance;

determines unresolved referents in the important discourse constituents;

determines potential resolving discourse constituents with potential to resolve referents; percolates relevance score of important discourse constituents through a reduced span of potential resolving discourse constituents; and

determines a reduced span of discourse constituents based on relevance score.

22 - 33. (canceled)

34. (original): The method of claim 8, further comprising the steps of: determining a combined relevance score based on the relevance scores and non-structural relevance scores and percolating the combined relevance score.

35. (original): The system of claim 18, the relevance circuit determines combined relevance score based on the relevance scores and non-structural relevance scores and percolating the combined relevance score.

36. (canceled)

37. (previously presented): A hybrid text summarization system comprising:
means for determining discourse constituents for a text;
means for determining a structural representation of discourse for the text;
means for determining relevance scores for discourse constituents;

means for percolating relevance scores based on the structural representation of discourse comprising the steps of:

means for each discourse constituent leaf node, determining the number of subordinated edges plus one;

means for determining a score based on the inverse of the number of subordinated edges +1;

means for each discourse constituent node, assigning the score of a child discourse constituent node to the parent discourse constituent node, if the score is less relevant;

means for any subordination discourse constituent node, assigning the score of the subordinated discourse constituent node to the subordinating discourse constituent node if the subordinated discourse constituent score is lower;

means for assigning the relevance scores of any coordination discourse constituent node to each child discourse constituent of the coordination if it is lower;

means for determining an adjusted relevance score based on the score and the subordination level; and

means for determining a hybrid text summary based on discourse constituents with relevance scores compared to a threshold relevance score.